DOCUMENT RESUME

ED 166 866 EC 112 505

AUTHOR	Lynch, William W.
TITLE	Guidelines for Applied Uses of Observation Instruments in Teacher Education. Final Report.
INSTITUTION	Indiana Univ., Bloomington. Center for Innovation in Teaching the Handicapped.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
PUB DATE	Jan 76
GRANT	OEG-9-242178-4149-032
NOTE	45p.
AVAILABLE FROM	Indiana University, Center for Innovation in Teaching the Handicapped, School of Education, Bloomington, Indiana 47401 ($1.50)
EDRS PRICE	MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS	*Classroom Observation Techniques; Cost Effectiveness; *Guidelines; *Handicapped Children; Higher Education; *Informal Assessment; Information Dissemination; Inservice Teacher Education; Reliability; Test Reliability; Test Validity; Validity

ABSTRACT

Guidelines for the use of observation instruments in pre-and in-service teachers of the handicapped are provided. It is explained that guidelines have been developed according to the following five principles: (1) the observation instrument must be relevant (including empirical, normative, and interpretative relevance); (2) methods of gathering observation data should be feasible and cost-effective (including collection, storage, and retrieval of the information); (3) observation data should be reliable (including both accuracy and dependability); (4) observation data should be valid according to three criteria - context, process, and outcome; and (5) observation data should be planned so as to provide maximum impact (which is the ultimate test of classroom observation methods). A checklist with questions to assess the instrument according to the five principles is included. (Author/PHR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED166866

GUIDELINES FOR APPLIED USES OF OBSERVATION INSTRUMENTS IN TEACHER EDUCATION[1]

William W. Lynch

January, 1976

Final Report 29.33

Center for Innovation in Teaching the Handicapped

Indiana University

[1]The writing of this paper was supported by grant #OEG 9-242178-4149-032 from the U. S. Office of Education, Bureau of Education for the Handicapped, to the Center for Innovation in Teaching the Handicapped. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

EC112505

PREFACE

This report deals with practical problems involved in choosing or designing instruments for observing teaching, as well as problems in using results from these instruments to help teachers (or prospective teachers) improve their skills. The guidelines that constitute this report have grown out of the author's experiences in studying the needs and instructional functions of teachers of the handicapped in a series of projects carried out at the Center for Innovation in Teaching the Handicapped (CITH) since 1969.

The report is intended to serve two purposes. First, most of the ideas set forth in the report are direct outcomes of CITH projects and, as such, should be reported to those sectors of officialdom that need such information. Second, the report is written for dissemination by CITH to its clientele of teacher educators and public school personnel who look for guidance in improving the effectiveness of school personnel working with the handicapped.

Most of the author's projects have been carried out under the general rubric, "The Cognitive Demand Skills of Teachers of the Handicapped." These have been efforts to identify and validate specific interactive skill patterns of teachers of the mildly handicapped, particularly those pertaining to the more cognitively oriented phases of academic instruction. The early projects centered around an observation system, the Individual Cognitive Demand Schedule (ICDS), developed by the author and Carole Ames (Lynch & Ames, 1971a, 1971b, 1971c, 1971d, 1972; Lynch, Ames, Barger, Frazer, Hillman, & Wisehart, 1973). The ICDS was used to record teacher-pupil interactions in Project PRIME, a large scale evaluation of "mainstreaming policies in the State of Texas."

As we gained experience in observing both regular and special classes in public schools, the limitations and advantages of observation methods for applied uses became increasingly clear. Observation methods provide powerful tools for the improvement of teaching, since teachers can be helped to see the details and patternings of classroom events. But observation methods can also be misleading, particularly when methods are used that are not relevant or valid for particular classroom conditions. Since the first rash of enthusiastic uses of observation methods in the late 1960's and early 1970's, the trend has been away from global methods toward more specialized methods with demonstrable relevance for specific instructional contexts and problems.

Realization of limitations of global instruments led the author to abandon efforts to develop a second, simplified and applied version of the ICDS. Extensive consultation with supervisors of student teachers and public school personnel, together with further experience observing in widely varying school contexts, led to the realization that publication of a simplified global instrument would only encourage applications of what Kaplan has termed "The Law of the Instrument,"--"Give a small boy a hammer, and he will find that everything he encounters needs pounding" (Kaplan, 1964, p. 28).

Practitioners in teacher education need to make good uses of observation methods. To do so, they need to be able to select or adapt existing instruments to fit the requirements of their situations. Frequently, they will need to devise their own instruments. It is hoped that this report will furnish guidance for such efforts.

The ideas in this report could not have reached their present state of articulation without the shared experience and wisdom of colleagues in the Center for Innovation in Teaching the Handicapped. Carole Ames, Mary Ella Brady, Robin Dalton, Louis Epstein, and Stephen Hillman have contributed substantially to our understanding of observation methodology by their efforts to interpret the complex events investigated in our projects. Melvyn Semmel and Merrill Sitko have provided interpretive insights and critical feedback that have opened new vistas. Ted Frick and Judy Agard, in their incisive analyses of Project PRIME observational data, have added importantly to progress in interpretation and analysis of observational data obtained under a wide variety of public school conditions. To these individuals and to students and staff members who have participated in many discussions at CITH, the author expresses gratitude. It should be understood, however, that none of these individuals shares the author's responsibility for any shortcomings in this report.

Guidelines for Applied Uses of Observation Instruments in Teacher Education

William W. Lynch
Center for Innovation in Teaching the Handicapped
Indiana University

This report sets forth some guidelines for the use of observation instruments in teacher education--techniques for systematically observing and coding various aspects of the teaching process. Emphasizing application rather than research, this report is addressed to teachers, teacher educators, and supervisory personnel whose primary interest in observation is the improvement of teaching in school and teacher-education settings. These guidelines supplement other published discussions of classroom observation methods mainly concerned with technical and scientific adequacy. The guidelines take into consideration the typical needs, problems, and resource limitations facing practitioners. It is hoped these guidelines will encourage the intelligent use of observation methods--uses consonant with the best technical and scientific knowledge that, at the same time, meet the demands of practical situations.

Since the early 1960's, there has been a very rapid growth in the use of observation instruments in research on teaching. These techniques have also had strong appeal for teacher educators, an appeal undoubtedly due in large part to their apparent "face validity." Observation techniques seem to deal uniquely with the "real world" of teaching. The numerous instruments that have been created and disseminated reflect a bewildering diversity of theoretical and practical concerns (Rosenshine & Furst, 1973: Simon & Boyer, 1967, 1970). Although most instruments

have been developed for research purposes, their appeal for practitioners has been strong. Frequently in American education, the development of new technical devices has been followed by an eagerness to use them in applied situations. Unfortunately, use of new techniques in educational practice is not always accompanied by a clear understanding of their range of usefulness or their limitations.

The guidelines discussed below pertain to all techniques calling for one or more observers to view and listen to a teaching situation (or an audio- or video-recording of a teaching situation) and then to encode what is perceived into a permanent, organized record for subsequent use in efforts to improve teaching. The methods include category systems, sign systems, and rating systems.[2] The guidelines do not include techniques dependent upon the observer's memory or general impressions rather than upon the systematic coding of classroom events. Also excluded are anecdotal records and other means of recording the events of teaching in a continuous narrative record (e.g., specimen records, participant-observer diaries, and other methods of behavioral ecology and "microethnography"). The guidelines are applicable both to existing published instruments and to the development and use of "homemade" instruments. The emphasis throughout is on principles for ensuring that methods are suitable for particular purposes, rather than on details of technique.[3]

---

[2]See Semmel and Thiagarajan (1973) for an excellent introduction to the characteristics of observation systems.

[3]Herbert and Attridge (1975) have presented 33 criteria for developers and users of observation systems and materials that emphasize desiderata for adequately communicating information about observation systems.

This report grows out of the experiences of the author and his colleagues at the Center for Innovation in Teaching the Handicapped (CITH at Indiana University. Over the past six years, the professional team at CITH has utilized observation methods in an effort to develop better ways of helping teachers of handicapped children improve their effectiveness. The guidelines are based primarily on experience with teachers of the handicapped, but examples have been expanded to represent a wide variety of settings and needs.

The guidelines are developed according to five general principles:

1. Observation methods should be chosen so as to be relevant to the concerns of those who use them.

2. Methods of gathering observation data should be feasible and cost-effective.

3. Observation data should be sufficiently reliable for the uses to be made of them.

4. Observation data should be valid for the situations observed and for the uses to be made of them.

5. Uses of observation data should be planned so as to have optimal impact on the teacher's competence.

I. Relevance

When an observation system is to be used to help a teacher improve some aspect of classroom practice, the process of conducting observations and the form taken by the resultant data should be easily and understandably related to the concerns, perceptions, and actions of the teacher. In educational research, it is customary to assume that a data-gathering technique is useable if it meets the technical criteria of reliability

and validity. These criteria are, of course, essential. But an observation system that meets conventional standards of reliability and validity may be quite useless if the information it yields bears little or no relationship to the problems of the classroom as seen by the teacher and whoever else is involved in applied uses of the data.

The criterion of relevance as treated here is related to the criterion of validity. Both concepts refer to the significance of what is measured by an instrument. The intent of this section is to encourage the potential user of an observation method to reflect carefully on the relationship of the technique to what the user knows about the situation and to what he considers to be important. The criterion of relevance emphasizes the importance of the realities of the situation as seen by those who are closest to it--a phenomenological emphasis. The criterion of validity emphasizes more formal, precise, and systematic relationships among variables--a scientific emphasis.

The general criterion of relevance has three aspects: empirical, interpretive, and normative. The empirical aspect of relevance addresses the question: "Is the observation instrument sensitive to the important facts of the teaching situation?" The interpretive aspect addresses the question: "Does the language of the instrument make sense to its users?" And the normative aspect is concerned with: "How does the instrument and its uses relate to the dominant values of the teaching situation?" Each aspect of relevance will be explained with some positive and negative examples.

Empirical Relevance

The practice of teaching encompasses an enormous variety of situations, tasks, materials, human characteristics, and modes of human action. Therefore, it is very important that the user of an observation method be certain that the method chosen fits the situation. It has become a very common experience for users of observation systems to discover that a given technique completely "misses" the significant features of a teaching situation. They discover that the instrument selected has little or no relevance for the facts in the situation at hand. Empirical relevance is fidelity to those facts that the user knows to be salient and important in the situation.

Example 1. A special class teacher of a primary EMR class has been regularly scheduling "show and tell" period at the beginning of each morning to provide an opportunity for each child to practice oral expression and to motivate the active verbal participation of each child in a semi-structured group activity. She is uncertain as to whether she is succeeding in encouraging each child to contribute; she has the uneasy feeling that some children are "hogging the stage." So she asks the teacher's aide assigned to her to observe "show and tell" for five successive days with a simple observation schedule that tallies: each time each child in the group speaks; whether or not the teacher has called on that child; and whether or not the teacher praises the child for his/her contribution. At the end of the five days, she discovers that two children out of the 15 in her class never participated and were never called on, while two other children contributed more than three times during each "show and tell" period. This simple, "homemade" observation system has empirical relevance because it focuses on a child

response that is germaine to the activity at hand, and it provides a record of relevant behaviors for each individual child.

Example 2. A supervisor of student teachers is concerned with whether or not his supervisees are keeping pupils actively involved during arithmetic instruction. He selects the Flanders Interaction Analysis Category System, an observation instrument that codes verbal interactions of teachers and students every three seconds into one of eleven categories. Eight of these categories represent different types of "teacher talk," two categorize "pupil talk," and the eleventh category represents "silence or confusion." Each time the supervisor observes one of his student teachers during arithmetic, the pupils are engaged in seatwork. Ninety percent of the codings fall in the category "silence or confusion." For that student teacher's class, the Flanders instrument was irrelevant simply because it was designed for an interactive type of teaching. Furthermore, it provided no record of individual pupil behavior. A simple "on-task--off-task" coding system that systematically scans the members of the class on a time sampling basis for recording each member's task involvement would have had more empirical relevance for this situation.

Interpretive Relevance

Teachers use many different kinds of concepts and interpretive models to make sense of classroom events. Some of these concepts may reflect their formal training, or the framework of a particular curriculum. As Joyce and Weil (1972), among others, have suggested, teachers can operate on the basis of alternative models. Classroom events, and the role of the teacher in influencing and responding to those events, can be interpreted by different models. A teacher who views interpersonal

influence from a Skinnerian point of view is likely to see things very differently from a teacher who has a Rogerian orientation. One teacher may have been trained in one or more formal, systematic models, while another teacher's frame of reference may be a highly individual one. Observational data that are expressed in the meanings of the teacher's frame of reference are interpretively relevant; those whose meanings are alien to the teacher are interpretively irrelevant.

Example 1. A teacher of a class for emotionally disturbed children has been trained in a behavior-modification approach that emphasizes systematic uses of social reinforcement in the management of disturbing behavior. Several children seem to be getting worse as time goes on. The teacher suspects that he may be unwittingly reinforcing certain behaviors. So he asks a fellow teacher to observe his class, focusing on selected disturbing behaviors of the two children and the teacher's own verbal and non-verbal behavior toward those children. He discovers that there is a relationship between disturbing behaviors (hitting neighbors and dropping materials) and the teacher's visual attending behaviors. He deliberately modifies his non-verbal reactions and verbally rewards the children for behaving more appropriately. The incidence of disturbing behavior diminishes. The observational data are completely consonant with the teacher's "behavior modification" orientation. The observation technique is interpretively relevant.

Example 2. A teacher of reading is concerned with whether or not he has been encouraging his pupils to utilize appropriate word-attack skills when they encounter new words during oral and silent reading. His supervisor has just acquired a new classroom observation system designed to code teacher-pupil interactions during reading instruction.

The supervisor observes the class with the instrument during reading period on two different days. Then the supervisor meets with the teacher to show him the results on a summary sheet. The instrument classifies all questions during reading into six categories--recognition, recall, translation, conjecture, explanation, and evaluation. It also gives a record of the "congruence" of pupil responses to teacher questions. After discussing the results, the teacher and his supervisor realize that the instrument provides no information relevant to pupil efforts to decode new words. Rather, it is designed to provide a record of teacher strategies in stimulating reading comprehension. They later discover another observation system--the Oral Reading Observation System--that is interpretively relevant to the teacher's interests.

Example 3. An intermediate EMR class teacher learned in his science methods course to classify instructional objectives and classroom activities according to Gagne's eight types of learning. As part of an inservice training program on improvement of teacher questioning skills, a tape recording of one of his science classes is analyzed with the Florida Taxonomy of Cognitive Behavior, an observation system based upon Bloom's Taxonomy of Educational Objectives--The Cognitive Domain. During the discussion of the analysis, the teacher objects that the observation record has no relation to what really went on or to his lesson plan. Even though both systems attempt to interpret the cognitive events of a classroom, they do so with different terminology and a different logic, hence creating a mismatch for the teacher.

Normative Relevance

The normative aspect of relevance refers to the teacher's values and goals. In virtually every aspect of teaching, the teacher is trying to maximize certain results in keeping with his/her values and goals for the situation. The goals reflect the teacher's own values as well as the values of the school and the particular class. When observation methods are used, the information that they produce will quite naturally be interpreted in the light of the value-related concerns of the teacher or other user. For example, one teacher may place a premium on having a quiet, orderly class. Another teacher may especially prize openness and spontaneity. Each teacher would probably react differently to an observation system that is designed to measure teacher effectiveness in classroom control.

Example 1. In a preservice training program for special education teachers, students participate in a semester-long practicum in which they teach or tutor mildly handicapped children in order to learn specific interactive teaching skills. Included in the training are learning systematic observation skills, practicing specific interactive skills, and obtaining feedback from observational records of their own performance as they teach. Early in the program, they undergo intensive training in "higher level cognitive demands" during which they learn the different cognitive levels of teacher questions and pupil responses. The students begin to be confused and demoralized, even though they are making some progress in identifying and producing higher and lower level "cognitive demands" as defined by the observation system in use. The instructor finds that students are devoting much too much time to such apparently routine matters as getting the children settled before starting a lesson in which

questioning skills are to be practiced. There are frequent complaints that the practicum is not preparing the trainees for student teaching. Students express concern over their inability to control and motivate children. Finally, the instructor realizes that the focus on cognitively-oriented interactions is not relevant (normatively) to the students' dominant concerns for management and motivation. After the program is revised to provide for initial mastery of a repertoire of management and affectively oriented rapport-building techniques, students feel more ready to consider the uses of questioning techniques, and the "cognitive demand" observation system then has normative relevance.

Example 2. A junior high school teacher of a home economics class for educationally handicapped children (EMR, LD) complains to her department chairman that too often students are not getting started on projects in clothing construction because they are not following simple directions. They decide together to gather observational data during the teacher's introduction to a new project to see if (1) the teacher's instructions are complete and clearly expressed and (2) whether the teacher is checking to verify each student's understanding of the directions for starting the project. They construct a check list of each key concept for a pattern cutting project and the essential operational steps that each student must take. After identifying students who have problems, they make an observation record containing the names of the students and the following categories: "Teacher asks probing question on concept X;" "Teacher asks student to state in order the first _____ steps;" "correct pupil response;" and "incorrect pupil response."

The class is observed by the chairman, and it is discovered that the teacher glossed over two essential steps. Three of the five problem students could not correctly state these steps when asked to do so. The observation schedule has been specifically designed to focus on well-defined concerns of the teacher. It has high normative relevance (as well as high empirical and interpretive relevance).

Once the prospective user of an observation instrument has a reasonably clear idea of what the characteristics of the situation to be observed are, how the information collected is to be construed, and what goals the information is to serve, then decisions can be made on the desired features of the instrument itself. These decisions include choosing the kinds of information to be recorded, the amount of detail and complexity desired, and the extent and kind of quantitative differentiation desired.

Any of five major kinds of information may be selected for observation and coding on a classroom observation instrument. They are teacher behavior, pupil or learner behavior, patterns and sequences (combinations of teacher and pupil behavior), context characteristics, instructional and physical resources of the teaching situation, and time. These kinds of information may be recorded in all possible combinations, depending upon need.

The amount of detail and complexity desired ranges from a focus on single, isolable behaviors (e.g., pupil hand raises) to subtle, covert phenomena to be inferred from combinations of behaviors available to the observer (e.g., judging divergent, inferential thinking by a pupil; or "observing" that a class climate is "warm" and "democratic").

The type of quantification derivable from coded observations range from simple "all-or-none" designations of categorized events, through simple counter frequencies, to proportions, ratios, or scale values based on some statistical rationale.

In summary, selection or design of an observation instrument should begin by a preliminary analysis of the situation to be observed, the kind of interpretation one wants to make of the data, and clarification of the purposes to which observation data are to be applied. Once having established this framework--a clarification of relevance--the user can determine what features of information the instrument should produce.

II. Feasibility

Since observation methods can be time consuming and cumbersome, it is important to design and plan them to take best advantage of available resources and to maximize their benefits. This section on feasibility suggests ways in which resources can be optimized in collecting, storing, and retrieving information. The resources considered are human resources, time, materials, and technical equipment.

Collection of Data

Once the focus and purpose of observation has been decoded, the first question in planning observation is "Who will observe?" In a training situation (e.g., a preservice teacher education program), the choice lies among the instructional personnel and the trainees themselves. In inservice programs, potential observers include supervisory personnel, paraprofessionals and aides, fellow teachers, pupils and (through the use of tape recordings) the observed teacher him/herself.

The choice of observer depends upon the kind of observation methodology selected, the training it requires, the time and sampling requirements necessary to achieve adequate sampling of situations and sufficiently reliable data, and how the observer might be affected by the experience of observing or being involved in any follow-up activities. Training procedures and the

time they take depend upon the simplicity or complexity of the events to be observed and the ease of coding and recording. Simple, easily isolable behaviors that must only be identified or counted (for example, counting how often a single pupil raises his hand during a lesson) require little training time and can be done by persons of varying backgrounds and abilities. Complex events or coding systems that require the observer to engage in difficult inferential and judgmental processes while data are being recorded require extensive training. Also, observers must be selected who are capable of learning the system and of maintaining a dependable performance.

Under some circumstances, it is desirable to select observers who will learn from the experience of observing because they are in the process of training for teaching roles. It has been shown that teachers-in-training have acquired teaching skills more easily when they have had the experience of systematically observing and coding others performing the same skills (e.g., Greenough, 1975). In inservice training, peer observation (observing another teacher) has the advantage of ensuring that the observer has a frame reference similar to that of the teacher being observed. Peer observation also tends to ensure that pre-planning and designing of observation procedures will be done with the needs of the teacher in mind rather than following a mechanical procedure that may not be fully relevant or well understood.

In many school situations, it turns out that the teacher is the best person to do the observation simply because he/she has the best opportunity to obtain appropriate samples. This typically requires the teacher to go over a tape recording, but, for some instructional activities, the teacher may be able to keep an on-going record of selected pupil be-

haviors, uses of instructional resources, or even the teacher's own actions. Tape recordings of teaching also make it possible for the teacher and a peer or consultant to code the recording together. (See Allen, Barnes, Reece, & Roberson, 1970, for suggestions to teachers for recording and observing themselves.)

Reciprocal peer observation of real or simulated teaching has advantages in both pre- and inservice teacher education. Trainees become sensitized to relevant features of the teaching situation and receive feedback on their own performance.

In some situations, pupils themselves can be used to gather observation data, as long as the process does not interfere with normal class participation and the method is within the pupils' capabilities. With check lists, simple diary records, and the preservation of written work, pupils can furnish the teacher with evidence of their own participation, the activities of other pupils, and the ways in which instructional materials are being used. Depending upon the maturity of pupils, it is also possible by gathering pupil reactions to gain insights into a teaching situation that no other observer could furnish (e.g., pupils' own perceptions, feelings, and non-observable learning processes).

Another valuable use of student observers is in group work where a pupil may be formally assigned the role of observer. In this role, the child keeps a systematic record of the transactions of a group. Although this type of observation is typically for the principal benefit of the group members themselves, the record may also furnish useful data for the teacher's own edification.

Instrument Design

The format for coding data should be designed for ease of coding and for ease of retrieval. The following suggestions are divided into two parts. The first part lists design considerations for pencil-and-paper coding. The second deals with the uses of technical devices for collecting observation data.

The principles that should govern the design of paper-and-pencil record forms for classroom observation are based upon considerations of (1) making the coding process as easy as possible; (2) ensuring that the information collected is sufficiently complete for the uses to be made of it; (3) ensuring that information can be summarized and used with as little labor as possible.

In designing a record for the collection of observation data, it is extremely important that the process of recording the information be as simple as possible. The observer must be left as free as possible to view and listen to the teaching as it proceeds. The page format needs to be amply spaced and the various category designations clearly identified so the observer does not have to spend time looking for information or finding his/her place. The process of marking should be as simple as possible. If jots or check marks can be used instead of writing down letters, numbers, or words, so much the better.

Only those items of information that the user is certain will be needed should be included. It is frequently tempting to add items to an observation system just to make it more logically complete or because some material "might be nice to have sometime." These temptations should be scrupulously avoided. On the other hand, one must be certain that the instrument is com-

plete. Some items of information that did not seem important at the time of observation may prove important later, especially such identifying information as: names, dates and time, place, type of instruction, type of group situation, physical arrangements, instructional materials and equipment. If time relationships are important (such as the rate at which events occur), the observer needs to have a clock or stop watch conveniently located so time intervals can be noted. If the sequence of events needs to be preserved, the format of the observation form should conveniently provide for this information. It is often useful to know who the participants were in events being recorded. When pupil behavior is important (as it normally is), then the record form and coding system should provide a convenient way of representing individual pupils by initials, code numbers, or check-mark spaces.

Retrievability of information from a record form is of greatest importance in applied uses. Whatever can be done to put information in a form that will not require laborious and time-consuming summarizing, counting, summing, etc., is effort well spent. If the data are to be summarized in numerical form, the format of the record form should be such that totals, subtotals, and proportions can be easily and directly taken from the form. Frequently, information that is presented in the form of a graph is very useful. If so, the placement of tally marks next to category designations can be done in such a way as to virtually create the graph as observation proceeds.

For some uses of observation data, the qualitative features of teaching are more important than the quantitative. If so, the form for coding the information should facilitate subsequent retrieval. For example, when one is looking for certain patterns or sequences of interaction, these can be so designated on a record form. The symbols used in coding can be arranged

in such a manner that patterns and sequences can be easily recognized when a summary is being prepared or when teacher and colleague are going over the record later.

The collection of observation data by "hand" with paper-and-pencil methods, no matter how well planned and designed, will always be somewhat cumbersome. If available, technical devices should be used to make the process more efficient.

Three types of equipment facilitate data coding procedures. First, there are those simple mechanical means of recording data from which the user can immediately read off results. Simple timers and counters are the most familiar examples. For classroom observation in which frequency data on only one event are needed, a simple digital counter may be an extremely valuable tool for the teacher or teacher educator. Similarly, the portable or pocket electronic calculator can be of assistance, with the added advantage of performing more than addition manipulation with incoming numerical data (numerically designating intensity of some event may be provided for with this device).

Next in complexity are mechanical or electronic data recorders that create a retrievable record of frequencies by category and, in some cases, time intervals. Such equipment greatly simplifies the observer's task because, instead of having to write down items and scan an observation sheet, the coder typically presses buttons on a panel, using a number system to represent the events coded. With relatively little practice, observers can use the button panel without having to look at it at all.

At the third level of sophistication in instrumentation is the inputting of observation into a computer system. The mechanics of coding are no different from those described in the preceding paragraph. The computer can

not only store information but also summarize and transform it for immediate retrieval and interpretation. For example, immediately after completing a lesson, a teacher can be furnished with a printout that meaningfully summarizes the lesson with tables, graphs, or even narrative interpretations. Or the teacher may be given immediate and continuous feedback on a television screen or auditorily through an earphone while teaching is going on. (See Semmel, Olson, & Weiske, 1972 for details of such technology as developed at the Center for Innovation in Teaching the Handicapped.)

III. Reliability

The user of data from an observation system needs to be certain that the data are sufficiently "accurate" or "dependable." The determination of the reliability of observation instruments and the calculation of reliability coefficients involve relatively complex technical questions that will not be broached in detail here. (See Frick and Semmel, 1973 and Rowley, 1975 for useful discussions.) But some general suggestions follow for ensuring that observation data are sufficiently reliable for the practical purposes the user has in mind.

"Reliability" has two important meanings when applied to observation data. First, it means the degree to which the method is dependably objective. Do two or more similarly oriented observers, using an instrument in a technically correct manner, obtain the same results? Poor reliability (in this sense of objectivity) primarily stems from two difficulties. One is the problem of variations in the meanings of categories from one observer to another. Ambiguity of definition tends to be a problem when categories require the observer to make inferences or to use evaluative judgments. The other problem is that observers may not focus their attention on the

same events or behaviors as they use the instrument. Classroom events are complex and fast moving, and it is not always easy for an observer to systematically focus attention on key elements in the situation.

In applied uses, two simple procedures can be followed to help correct problems of definition and attentional focus. The first is to have two observers (a teacher and a colleague, for example) present to each other examples for each category or term in the instrument, and then decide whether or not there is mutual agreement in defining the categories. The second and more thorough procedure is to have two or more observers record the same classroom events and then compare the results for agreement by each event, discussing and clarifying differences in meaning or data when the coding of a given event resulted in disagreement.

A refined and particularly powerful application of the latter procedure for developing interobserver agreement has been developed at the Center for Innovation in Teaching the Handicapped. This procedure, called DITRMA, provides for several observers to code their observations of videotaped classroom events on button boxes (similar to touchtone telephone panels) and receive continuous, immediate feedback from a computer on whether or not there is observer agreement. Whenever observers disagree, the videotape stops immediately, and the observers receive information from the computer on how each observer coded the event. Then they must discuss the problem and recode with unanimity before the videotape record resumes. This system has proven very efficient in training observers to satisfactory levels of agreement, typically cutting in half the observer training time needed for most instruments.

If there is no opportunity for an observer to check his record against that of an equivalently trained observer, then the observer needs to determine

the consistency of his own coding (intra-observer reliability). This is a common situation in applied settings. The procedure is best done with tape recordings. One should obtain a good sample of classroom behavior on video- or audiotape and code it completely once. Then, preferably after an intervening time period of at least several hours, recode the same tape. (See Frick and Semmel, 1974, for methods of computing reliabilities.) Such a check typically reveals that not only does an observer tend to perceive differently events in the same situation, but that meanings of categories have an unfortunate tendency to vary. These problems arise partly because of the similarities between many classroom events, and partly because immediately surrounding events may cause the observer to construe an event in different ways. Such a rechecking of his/her own observations is constructive for sharpening the observer's definitions, as well as for training in more disciplined methods of observation.

The second aspect of reliability is the representativeness of a given sample of observation data when compared with a fuller record obtained over a longer period of time. In other words, can the user depend upon a given sample as typifying the situation being observed? This can be a serious problem in classroom assessment. For example, if an observer (e.g., a supervisor) codes 20 minutes of a classroom discussion and tells the teacher that there were too many "lower level questions," would he have arrived at the same evaluation if he had observed for an additional 20 minutes or had observed on another day?

Research on classroom behavior suggests that a single short sample obtained on an observation system frequently may be too undependable to serve as a base for advice or other kinds of feedback to teachers (Brophy & Good, 1974; Lynch & Ames, 1971b). For example, an hour's sample of class-

room behavior may, when compared to 20 hours of observation, turn out to be very unrepresentative. Rowley (1975) has demonstrated that accumulating several relatively short samples from the same situation (say, over several days) results in higher reliabilities than devoting the equivalent amount of time to one extended observation. Of course, in most applied situations it is not feasible to keep gathering data until one is certain that the data are stable enough to be taken as typical. Usually the practitioner does not have the resources or technical skills to institute sophisticated methods of sampling in order to attain high levels of stability. Under these limiting circumstances, it is important to try to identify any factors surrounding observation data that may cause one to believe the data are not representative of longer samples. Probably the best source for such a check is the teacher of the class observer who may be able to identify circumstances that make an obtained sample of observational data unrepresentative. Factors that commonly make samples unrepresentative include: the presence of a stranger as observer; an unexpected or disturbing event that makes students more uneasy or distractible; unusual features of curriculum materials; and unusual behavior of individual pupils that affects others.

It is also important to recognize that the different stages of classroom activities will determine the frequency and intensity of events recorded through observation. Usage of certain instructional materials, for example, may predispose a class to certain kinds of classroom behavior that would not be typical if other materials were used. Furthermore, instruction goes through phases. For example, a teacher who is introducing a lesson may devote a relatively long period to eliciting child discussion in order to establish a frame of reference ("advance organizers") or to review past

work. The interactive behaviors for this phase might be very different from those of a later stage during which children are beginning to explore and utilize new concepts. If one were interested in the cognitive levels of discussion, the first phase might contain a large proportion of lower level questions. Then, only if the first phase had been successful might a teacher be able to elicit higher cognitive levels of pupil participation in subsequent phases. Observation data must properly reflect such stages. It is important that an observer using an instrument be aware of the structure of the teaching so as to avoid obtaining biased samples of classroom events.

IV. Validity

Curiously, the published literature on observation methodology devotes little attention to the question of validity. Perhaps this is because of the apparent "face validity" of most such methods. Nevertheless, the question of validity, commonly expressed as "does the instrument really measure what it is intended to measure?", is most important for judging the suitability of an observation instrument.

Part of the question of validity is met when the previously discussed questions of relevance and reliability have been answered. An observation system that meets the user's standards of empirical relevance, interpretability, normative relevance, objectivity, and representativeness is well on its way to being valid for the user. But the user also must consider other important information central to the user's concerns.

Specific meanings of the term validity, and the data-gathering and statistical procedures for estimating validity of instruments of measurement, vary considerably among researchers and measurement specialists, depending

upon the practical or theoretical interests of the investigator. When tests are developed for such practical purposes as evaluation or selection in education, "criterion-related" validity is given great emphasis. High correlations between scores on the test and some criterion measure are considered adequate evidence of either concurrent or predictive validity. Questions as to why the relationship is found are not considered important. On the other hand, when a researcher is interested in understanding the significance of test data with reference to some larger (usually theoretical) set of concerns, then the emphasis is on demonstrating a more complete set of systematic relationships between the instrument's measurements of a number of other phenomena. This is the concern for construct validity. Construct validity is established by showing relationships which verify a set of logically interrelated propositions. Rather than just showing one can predict some criterion with the instrument, construct validity consists of a set of relationships that clarifies how the instrument's measures fit into some larger framework. For example, construct validity might be demonstrated by a matrix of positive, negative, and zero correlations between scores on the instrument and other variables, each of which makes good theoretical sense ("convergent and discriminant validity"). (See Borich and Malitz, 1975, for an example of this approach to classroom observation instruments.)

At first glance, these questions of criterion-related and construct validity may seem very remote from the concerns of a practitioner who is using observation methods for the improvement of teaching. But these questions lie at the very heart of effective uses of observation. Unless it is clear that what is being observed relates in some demonstrable and understandable way to other important facts, then there is very little

point in using such a method. Of course, the practitioner normally has neither the time nor resources for carrying out systematic validity studies. But the practitioner does have a frame of reference and a set of standards for interpreting the teaching situations that are of concern to him/her. In a sense, this frame of reference and set of standards can serve as the equivalent of the scientist's theory in determining construct validity and the test technician's empirical criteria for criterion-related validity. The practitioner also has the opportunity to make rough empirical checks so as to ascertain whether or not the data obtained by an observation method appear to relate to other facts in sensible ways. In other words, to check on validity the user first should make as explicit as possible the reasons why the observation data may be significant by trying to specify cause and effect relationships between what is observed and other important features of the situation. Second, the user should take care to examine relations between observed data and other facts likely to indicate that what is being observed is meaningful. It is never wise simply to assume that an instrument is valid on the basis of its name or reputation.

Three sets of criteria can be considered in checking on the validity of an instrument for the applied situation: context criteria, process criteria, and outcome criteria.

Context Criteria

Context criteria answer the question, "How would the results of this observation method differ when used in a different situation (context)?" The important situational variations to be considered are mainly those of the teacher's own classroom--its curricular or program variations. If the user of the instrument can envision that the results of observation would vary in a predictable way from one type of instructional group or curricular

situation to another, then he/she made a good start in inquiring into validity. Being able to provide actual evidence to support the prediction is better assurance.

Examples. (1) An observation instrument that records media uses (chalkboard, overhead projector, tape recorder, etc.) in classroom instruction should yield predictable differences between classroom activities in which media presentations are expected to vary. For example, a science class requires many diagrams and pictures, while language arts relies on books, worksheets, and a classroom discussion.

(2) An observation system on affective communication should record much higher frequencies of verbal and non-verbal expression of feeling during a group counseling session than during an arithmetic lesson.

(3) An observation system that records cognitive levels of teacher questions and pupil responses should yield a greater proportion of high-level questions in a class session on solving mathematical problems than in a drill on addition facts.

Process Criteria

Process-related criteria serve to answer the question, "Is the observation system sensitive to the processes that are occurring in the classroom?" By processes is meant the patterned, sequenced interplay among teacher actions, pupil actions, and the instructional resources of the classroom. Most common, recurring classroom activities have what Barker (1968) and others call a "standing pattern of behavior"--a characteristic set of behaviors typical for each activity. Each activity calls for a pattern of behavior that occurs regardless of who the participants are. In addition to its "standing pattern of behavior," each activity also follows a typical sequence. The sequence may occur in a short period of time, as, for example,

during the typical teacher-question/pupil-answer/teacher-reaction sequence of the conventional class recitation. Or there may be longer sequences. For example, a typical lesson may consist of the sequence: motivating introduction, presentation of new information and rules, discussion by examples and applications, and a summary or review of main points to conclude the lesson. There also may be shorter sequences within longer sequences. And sometimes the sequential features of a class activity may involve a pattern of participation by different pupils, such as when an incorrect pupil response in a discussion is followed by the teacher's asking the same question of a brighter child who normally responds correctly.

Practical assessment of the validity of an observation system by using process criteria consists of two steps. First, the user identifies a classroom situation for which the observation technique might be used and examines the instrument to ascertain whether or not the resulting data would include information that would reflect those processes known to exist. If there are gaps in the instrument (events that should be recorded but are not provided for) or if there are many categories in the instrument that the user is certain will not occur, then the validity of the instrument is suspect and an alternative or revised instrument must be considered. Second, if the instrument appears to be relevant, then two or more classroom activities known to differ in constituent processes should be observed, just to be sure that the instrument is sensitive to the obvious process variations.

Of particular importance in judging the validity of an observation instrument against process criteria is making certain that the learner behavior is recorded in forms that well represent on-going actions of learners that have significance for the activity at hand. In the past it

has been popular to make teacher behavior the dominant focus of classroom observation instruments. While teachers can learn much from a more detailed picture of their actual teaching patterns, knowledge of pupil reactions is essential to improvement of classroom technique. It is difficult for a teacher to be fully aware of all on-going processes of the classroom. Information from systematic observation can alert the teacher to subtle and complex features of pupil behavior that are impossible to discern while teaching is going on. By receiving this information, teachers can learn to become more aware of those essential cues from pupil behavior on which they can base adaptations of their own techniques and modes of interpersonal influence. For example, if an observation instrument concerned with classroom management of deviant behavior is being used to help a beginning teacher improve his/her skill, the most important information to be provided is a detailed record of the immediate effects the teacher has on pupil behavior. In many situations, it is important that the observation instrument provide an accurate record of the behavior of each individual pupil, so that the teacher can know the differential effects of her/his actions on individuals rather than on simply some group average.

Product Criteria

Product criteria answer the question, "Do the observation data predict significant outcomes of classroom activities?" Since researchers have found it exceedingly difficult to establish a relationship between process variables and outcome measures, it may seem strange to ask practitioners to inquire into this tricky relationship. But process-product relationships cannot be ignored, and there are practical ways in which a teacher or teacher educator can analyze the relationship to be more certain that observation methods are valid.

In the first place, there is the question of the "theoretical" basis for a relationship between what an instrument measures and important teaching outcomes. In the practical situation the user can ask, "Is it reasonable to suppose that what we are observing is related to learning outcomes that can reasonably be expected from this teaching situation?" Or, "Can we think of a pattern of data obtainable through this instrument that would be evidence of teaching so outstanding that pupil gains would likely result?" These are purely speculative questions, but, if the answer is clearly negative or indeterminant, perhaps there is reason to further examine the content and procedures of the instrument to ascertain whether its predictive validity for product criteria might be improved. There are some observation instruments for which predictive validity or outcome criteria have been established (e.g., Brophy & Good, 1974; Dunkin & Biddle, 1974). Such instruments should be considered if they are relevant to the user's situation and if they meet the criteria of feasibility and impact (discussed below). At the same time, it should be recognized that some of the verified process-product relationships obtained with such instruments are quite context-specific. That is, teacher-pupil behaviors recorded on an instrument may be related to favorable outcomes only with certain pupil populations, in certain school contexts, or with certain teachers (e.g., Brophy & Evertson, 1975).

V. Impact

The ultimate test of classroom observation methods in pre- or inservice teacher education must be the pragmatic test. Do they improve teaching? This section presents guidelines for deciding what kind of impact observation data should have, as well as guidelines for choosing the kind of data and the

format for presenting the results of observational data gathering so the desired impact may be made.

Observation methods can be used in three major ways in both pre- and inservice teacher education. First, teachers (or prospective teachers) can use observation instruments themselves, either to observe other teachers or to observe themselves from videotape records. Observing while not actively teaching gives the teacher a chance to gain a new perspective. Guided observation can help the teacher become more sensitive to features of the learning situation that previously had been unrecognized or peripheral. Observing another teacher with an instrument can help a teacher learn how to execute a new technique. The experience of observing another teacher can also teach a teacher how to better anticipate the consequences of different events, including the teacher's own actions.

A second use of observation methods in teacher education is to present observational data to the teacher in order to foster improved interpretations of classroom events. Observational data can be a very helpful source of examples and illustrative patterns in helping teachers to understand new concepts, new methods, or new techniques. For example, a teacher who is using class discussion to carry out a "discovery method" of teaching science may gain a better understanding of the desired sequences of questions and problem-solving activities of such a method by seeing an observation record summary of a model lesson.

A third major use of observation methods is giving feedback to teachers on their own teaching. By presenting him/her with the results of observation, a teacher is given a focused and informative basis on which to base changes in efforts to improve. Feedback is probably the major use of observation instruments in teacher education.

The teacher trainer who wishes to use observation instruments or data from them in any of these ways must have a reasonably clear idea of how this use should affect the teacher or prospective teacher. Teaching can be described as the carrying out of some set of rather complex social-role functions. While we still have much to discover about how people learn to carry out these functions, it is apparent that a person's competence for any particular teaching function may be determined by any or all of the following: (1) the teacher's perceptions of the events in the teaching situation; (2) the teacher's concepts and beliefs about the situation; (3) the teacher's feelings and attitudes; (4) the teacher's goals and interests; (5) the teacher's thought processes; and (6) the teacher's overt and interactive performance. In deciding what impact observation data or methods are to have in a teacher education program or activity, the trainer-educator has to decide which of these aspects of the teacher's competence should be most affected. The following examples illustrate possible outcomes that could be the result of using observation:

A. The teacher's perception of classroom events (noticing and anticipating):

---greater awareness of selected on- and off-task behaviors of individual children
---greater sensitivity to verbal and non-verbal expressions of feelings by pupils
---ability to discriminate different ways in which children are using materials in an open classroom arrangement
---greater awareness of the amount of time a teacher allows different children to respond to the teacher's questions during a social studies recitation.

B. The teacher's hypotheses, concepts, and beliefs about classroom events:

---learning the difference between inductive and deductive questioning sequences
---understanding the functions of probing as a response to pupil contributions in a discussion group activity
---seeing the relation between pupil differences in readiness and pupil need for structure and guidance in a skill learning activity
---discovering which types of self-instructional materials work best with pupils at different stages of learning.

C. The teacher's feelings and attitudes:

---becoming more acceptant of a pupil towards whom the teacher feels antipathy
---learning to be less apprehensive about an aspect of teaching that has been highly anxiety producing (e.g., a beginning teacher's worry that a 'free play' period might get out of hand)
---becoming more open and flexible toward using varied alternatives to a textbook-workbook curricular format
---learning to control one's anger at acting-out, aggressive pupil behavior.

D. The teacher's intentions and goals:

---formulating realistic daily outcomes for target children in reading
---shifting from a set to "cover the material" in a lesson to one of ensuring every pupil a chance to demonstrate mastery of performance objectives
---combining purposes for an activity (e.g., emphasizing "appreciation" as well as "understanding" in poetry reading)
---giving a shy child opportunities to participate successfully
---providing students more opportunities to apply rules and generalizations in a variety of practical situations.

E. The teacher's thought, decision-making, and problem-solving processes:

---diagnosing individual learning problems
---choosing materials or classroom teaching aids to fit individual child needs
---figuring out how to encourage members of a class to take on more responsibility for their own classroom management
---deciding on what kind of prompting strategy to use when a child makes a miscue in oral reading.

F. The teacher's performance:

---distributing the teacher's attending behaviors (eye gaze, eye contact) more equitably to members of a group
---keeping voice volume high enough to be heard in the back of the room
---selective use of a reinforcing behavior with target children
---sequencing questions from concrete, recall questions to inferring and problem solving in a science lesson
---adaptive prompting of oral reading miscues.

In deciding what aspect of the teacher's competence to emphasize by using observation in training, a general rule is to choose the observation information and the format that have the greatest chance of involving the

teacher directly in the relevant aspect(s). In other words, if the impact is to be on teacher perception, the observation instrument or data should rather directly and immediately involve the teacher in noticing. If a change in teacher conceptualization of teaching events is sought, then the observation data should encourage the teacher to conceptualize. If feelings or attitudes are the impact target, the observation data should be capable of eliciting feelings or attitudes in productive ways.

After one is reasonably certain as to what changes in the teacher are most important, then a choice must be made regarding the best format for representing observation data. Figure 1 summarizes the principal types of information gathered by observation methods and the various ways in which that information may be represented or symbolized. A few examples of particular formats for feedback information have been entered in the matrix of Figure 1 as illustrations. (The reader should try to fill in other cells to test his/her understanding of the options possible.)

In order to make a choice of format for communicating observation data to teachers, one must have some feeling for what the process of change for the teacher is likely to involve. This requires a third decision, namely, a decision as to what the conditions should be under which observation data are to be given. Some common options are:

--direct feedback during teaching

--direct feedback immediately after teaching

--feedback coupled with a tape recording

--comparison of observed events with a set of norms or a model

--comparison of an observation record with that of an earlier record of the same teacher

--an interpretative discussion of observation data with a fellow teacher or a consultant

--an evaluative summary from a supervisor or consultant (e.g., "These are the things you did well...; you need to improve on the following...")

In general, decisions as to target changes sought, format, and conditions of use require some prior understanding of the nature of the problems being confronted by the teacher, as well as the teacher's capabilities for bringing about constructive change. At the present stage of knowledge in teacher education, these are matters of judgment and a certain amount of trial and error. In general, it is important to distinguish between those phases of teaching that require the immediate perception of on-going events coupled with production of influence behaviors, and those phases that require careful conceptualization and systematic decision making. The first is a matter of developing habits; the second, a matter of thoughtfully directed acts. In the former case, observation feedback should be aimed at sensitizing the teacher to subtle cues, reinforcing the teacher's ability to anticipate pupil behavior, and reinforcing effective influence behaviors. For these training needs, immediate feedback is probably important.

Sometime, the goal may be to foster more thoughtful decision making by improving the teacher's ability to understand the cause-effect relationships of classroom events or the relationship of knowledge of pupil characteristics to classroom processes. In such cases, the format and the conditions of feedback should be planned so as to foster clarity of conceptualization by emphasizing the teacher's own concepts, language, and decision processes.

TYPE OF INFORMATION	RAW DATA	NUMERICAL SUMMARY (Frequencies; Ratios; etc.)	GRAPH	VERBAL RESUME	COMPARISON (w. norms; models; past records)	VERBAL EVALUATION
TEACHER BEHAVIOR	Tallies for no. of praise remarks to target pupils	Ratio of indirect to direct teacher talk (Flanders ID Ratio)		Narrative resume of types of 'desist' techniques used in lesson		
PUPIL BEHAVIOR	Tallies for no. hand raises				Graph showing ind. off-task behavior w. average for class by 5-minute blocks	
PATTERNS; SEQUENCES			Line graph of ratio of correct pupil responses following teacher question 5-minute blocks			"Your higher level questions are too often followed by irrelevant pupil replies."
CONTEXT			Bar graph of small & large group instruction			
INSTRUCTIONAL MATERIALS		Percentage time of use of equipment				

Figure 1

Format Options for Feedback

Summary

The following checklist is offered as a summary of the guidelines. The teacher educator, supervisor, consultant, or administrator who is considering an observation instrument as a means for helping teachers (or prospective teachers) improve their skills and understanding of classroom events should consider each point in evaluating the instrument's suitability for the training tasks at hand. The checklist is a series of questions pertaining to each of the five general principles around which the guidelines have been developed. A "no" answer to any pertinent question suggests there may be some problems with the instrument or its uses that should be corrected. Corrections to be considered include: selecting a new instrument; modifying the instrument under consideration; constructing a more suitable instrument; and modifying the data-collecting, retrieval, and/or interpretation processes. In general, the guidelines are written to encourage creative and constructive practices and to discourage uncritical, mechanical uses of observation methods in teacher education.

A Checklist

I. Is the instrument relevant?

1. Does the instrument record the events and features of the teaching situation that you know to be important and that you want to record?
   - ...teacher behaviors?
   - ...student behaviors?
   - ...group characteristics?
   - ...activity or task characteristics?
   - ...time?
   - ...sequences and patterns?
   - ...physical arrangements?
   - ...instructional materials and resources?
   - ...the identity of students?

2. Is the instrument free of extraneous and trivial information?

3. Do the instrument's categories and language make sense to you and to the students with whom you plan to use it?

4. Is the information yielded by the instrument clearly related to the important concerns and training needs of the teachers with whom the information will be used?

II. Are the procedures called for by the instrument feasible for the resources available?

A. Personnel

1. Have you identified persons to serve as observers who will be available and willing when needed?

2. Is each person who is to collect data with the instrument qualified and adequately trained to use the instrument correctly?

3. Have you provided effective procedures for training observers?

B. Format

4. Is the format of the instrument designed for maximum ease of recording?

5. Is the instrument free of extraneous information and distracting, unnecessary detail?

6. Are data easily retrieved and summarized from the record forms of the instrument?

C. Equipment

7. Have you considered available types of equipment that may make coding easier and retrieval speedier and more convenient?

...counters
...calculators
...recorders
...computer-linked equipment

III. Is the instrument reliable?

1. Have you ascertained that observers using the instrument have been (or can be) trained to read a satisfactory level of agreement with a criterion (or expert observer) in coding the same classroom situations?

2. Have you ascertained that the samples of data to be collected with the instrument are representative of what would be found with more extensive sampling of the situations in which you are interested?

3. Have you distinguished between types of classroom activities or phases of classroom activities that should be sampled in order ensure that data are representative?

4. Have you taken into account factors in the situation to be observed that might make the data unrepresentative (e.g., the observer's presence, unusual happenings, etc.)?

IV. Is the instrument valid?

1. Can you demonstrate that the instrument differentiates among situations (contexts) in ways that you (and your trainees) can interpret meaningfully?

2. Can you establish that the data collected with the instrument are sensitive to various processes that you (and trainees) know to be important in the teaching situation?

3. Can you demonstrate (empirically or logically) that the data collected with the instrument are predictive of important pupil outcomes?

V. Does the instrument yield information that is likely to have beneficial impact on teachers?

1. (If relevant) Are you certain that data gathered with the instrument (or the experience of using the instrument) will enhance the teacher's awareness and perceptual sensitivity to important features of classroom events?

2. (If relevant) Will the data gathered by the instrument improve the teacher's understanding of (conceptualization of) classroom events?

3. (If relevant) Will the data from the instrument have beneficial effects on the teacher's feelings and attitudes?

4. (If relevant) Will the data from the instrument help the teacher to set goals or anticipate events more effectively?

5. (If relevant) Will the information for the instrument improve the teacher's skillfulness in direct, face-to-face, interactive teaching?

REFERENCES

Allen, P., Barnes, W., Reece, J., & Reberson, E. Teacher self-appraisal: A way of looking over your own shoulder. Worthington, Ohio: Charles A. Jones Publishing Co., 1970.

Barker, R. G. Ecological psychology: Concepts and methods for studying the environment of human behavior. Stanford, California: Stanford University Press, 1968.

Borich, G. D. & Malitz, D. Convergent and discriminant validation of three classroom observation systems: A proposed model. Journal of Educational Psychology, 1975, 67, 439-444.

Brophy, J. & Good, T. Teacher-student relationships. New York: Holt, Rinehart and Winston, 1974.

Brophy, J. & Evertson, C. Learning from teaching--a developmental perspective. Boston: Allyn and Bacon, 1975.

Dunkin, M. & Biddle, B. The study of teaching. New York: Holt, Rinehart and Winston, 1974.

Frick, T. & Semmel, M. Observation records: Reliabilities and observer agreement. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, April, 1974.

Greenough, K. The effects of anticipating or coding videotaped and written materials on developing the questioning skills of special education teacher trainees. Bloomington, Ind.: Center for Innovation in Teaching the Handicapped, Indiana University, Contract No. OEG-9-242178-4149-032, U.S.O.E., Bureau of Education for the Handicapped, Final Report 35.32, 1976.

Herbert, J. & Attridge, C. A guide for developers and users of observation systems and manuals. American Educational Research Journal, 1975, 12(1), 1-20.

Joyce, B. & Weil, M. Models of teaching. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Kaplan, A. The conduct of inquiry. San Francisco: Chandler Publishing Co., 1964.

Lynch, W. W. & Ames, C. Individual cognitive demand schedule. Bloomington, Ind.: Center for Innovation in Teaching the Handicapped, Indiana University, Contract No. OEG-9-242178-4149-032, U.S.O.E., Bureau of Education for the Handicapped, Technical Report 4.2, 1971a.

Lynch, W. W. & Ames, C. Differences among special class teachers in individualization of instruction. Bloomington, Ind.: Center for Innovation in Teaching the Handicapped, Indiana University, Contract No. OEG-9-242178-4149-032, U.S.O.E., Bureau of Education for the Handicapped, Final Report 4.3, 1971b.

Lynch, W. W. & Ames, C. Two experiments on modifying "cognitive demand" styles of elementary student teachers. Bloomington, Ind.: Center for Innovation in Teaching the Handicapped, Indiana University, Contract No. OEG-9-242178-4149-032, U.S.O.E., Bureau of Education for the Handicapped, Technical Report 29.2, 1971c.

Lynch, W. W. & Ames, C. Individual cognitive demand schedule: Observer's training manual. Bloomington, Ind.: Center for Innovation in Teaching the Handicapped, Indiana University, Contract No. OEG-9-242178-4149-032, U.S.O.E., Bureau of Education for the Handicapped, Technical Report 29.3, 1971d.

Lynch, W. W. & Ames, C. A comparison of teachers' cognitive demands in special EMR and regular elementary classes. Bloomington, Ind.: Center for Innovation in Teaching the Handicapped, Indiana University, Contract No. OEG-9-242178-4149-032, U.S.O.E., Bureau of Education for the Handicapped, Final Report 28.3, 1972.

Lynch, W. W., Ames, C., Barger, C., Frazer, W., Hillman, S., & Wisehart, S. Effects of teachers' cognitive demand styles on pupil learning. Bloomington, Ind.: Center for Innovation in Teaching the Handicapped, Indiana University, Contract No. OEG-9-242178-4149-032, U.S.O.E., Bureau of Education for the Handicapped, Final Report 30.3, 1973.

Rosenshine, B. & Furst, N. The use of direct observation to study teaching. In Travers, R. M. W. (Ed.), Second Handbook of Research on Teaching. Chicago: Rand McNally & Co., 1973, p. 122-183.

Rowley, G. A rationale for assessing the reliability of an observational measure. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C., April, 1975.

Semmel, M. & Thiagarajan, S. Observation systems and the special education teacher. Focus on Exceptional Children, 1973, 5(7), 1-12.

Semmel, M., Olson, J., & Weiske, W. An information and technical manual for the computer-assisted teacher training system (CATTS). Bloomington, Ind.: Center for Innovation in Teaching the Handicapped, Working paper 7.1, June, 1972.

Simon, A. & Boyer, E. G. (Eds.) Mirrors for behavior: An anthology of classroom observation instruments, V. 1-6. Philadelphia: Research for Better Schools, 1967.

Simon, A. & Boyer, E. G. (Eds.) Mirrors for behavior: An anthology of classroom observation instruments, V. 7-14 and Summary. Philadelphia: Research for Better Schools, 1970.